

# *A consistent interpretation of the stochastic version of the Ensemble Kalman Filter*

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

van Leeuwen, P. J. (2020) A consistent interpretation of the stochastic version of the Ensemble Kalman Filter. Quarterly Journal of the Royal Meteorological Society, 146 (731). pp. 2815-2825. ISSN 1477-870X doi: <https://doi.org/10.1002/qj.3819> Available at <https://centaur.reading.ac.uk/90781/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

Published version at: <http://dx.doi.org/10.1002/qj.3819>

To link to this article DOI: <http://dx.doi.org/10.1002/qj.3819>

Publisher: Royal Meteorological Society

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

## RESEARCH ARTICLE

# A consistent interpretation of the stochastic version of the Ensemble Kalman Filter

Peter Jan van Leeuwen<sup>1,2</sup> <sup>1</sup>Department of Atmospheric Science,  
Colorado State University, Fort Collins,  
Colorado<sup>2</sup>Department of Meteorology, University  
of Reading, Reading, UK**Correspondence**

P.J. van Leeuwen, Department of  
Atmospheric Science, Colorado State  
University, Fort Collins, CO 80526, USA.  
Email: peter.vanleeuwen@colostate.edu

**Funding information**

European Research Council via the  
CUNDA project, Grant/Award Number:  
694509

**Abstract**

Ensemble Kalman Filters are used extensively in all geoscience areas. Often a stochastic variant is used, in which each ensemble member is updated via the Kalman Filter equation with an extra perturbation in the innovation. These perturbations are essential for the correct ensemble spread in a stochastic Ensemble Kalman Filter, and are applied either to the observations or to the modelled observations. This paper investigates if there is a preference for either of these two perturbation methods. Both versions lead to the same posterior mean and covariance when the prior and the likelihood are Gaussian in the state. However, ensemble verification methods, Bayes' Theorem and the Best Linear Unbiased Estimate (BLUE) suggest that one should perturb the modelled observations. Furthermore, it is known that in non-Gaussian settings the perturbed modelled observation scheme is preferred, illustrated here for a skewed likelihood. Existing reasons for the perturbed observation scheme are shown to be incorrect, and no new arguments in favour of that scheme have been found. Finally, a new and consistent derivation and interpretation of the stochastic version of the EnKF equations is derived based on perturbing modelled observations. It is argued that these results have direct consequences for (iterative) Ensemble Kalman Filters and Smoothers, including “perturbed observation” 3D- and 4D-Vars, both in terms of internal consistency and implementation.

**KEYWORDS**

data assimilation, ensemble, nonlinear, particle filter, synchronisation.

## 1 | INTRODUCTION

A major breakthrough in data assimilation and Bayesian Inference for high-dimensional systems was the introduction of the Ensemble Kalman Filter (EnKF) by Evensen

(1994). The scheme uses an ensemble representation of the probability density function (pdf) of the state of the system and the propagation of the pdf in time is represented by the propagation of the ensemble members with the full model equations. The great practical advantage

over schemes like the Kalman filter and many variational methods like 4D-Var is that the model and the observation operators do not have to be linearised.

The EnKF update equation can be seen as an approximation to the Kalman Filter update equation. At observation times the Kalman filter assumes that the prior pdf is Gaussian, and also that the likelihood is Gaussian in the state  $\mathbf{x}$ . In that case the Kalman filter updates the mean of the prior Gaussian as:

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\mathbf{x}^f) \quad (1)$$

in which  $\mathbf{x}^f$  is the forecast mean,  $\mathbf{x}^a$  denotes the analysis mean,  $\mathbf{K}$  is the Kalman gain,  $\mathbf{y}^o$  is the observation vector, and  $\mathbf{H}$  is the observation operator that maps model states to observation space. *The notation used is  $\mathbf{y}^o$  for the actual observation, and  $\mathbf{y}$  for the case when the observations are seen as random variables. This distinction is of fundamental importance in this paper.* The Kalman gain is given by  $\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}$  in which  $\mathbf{B}$  is the prior state covariance and  $\mathbf{R}$  is the observation error covariance. The state covariance is updated in the standard Kalman filter as  $\mathbf{P} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}$ . For future reference, we also show the symmetric Joseph form of the covariance update as

$$\mathbf{P} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}(\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{R}\mathbf{K}^T, \quad (2)$$

which can easily be derived from the standard form by adding and subtracting the symmetric matrix  $\mathbf{B}\mathbf{H}^T\mathbf{K}^T$ , substituting for the additive term  $\mathbf{B}\mathbf{H}^T\mathbf{K}^T = \mathbf{K}(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})\mathbf{K}^T$  and collecting terms.

The original Ensemble Kalman Filter (EnKF) in Evensen (1994) implements the Kalman filter update equation on each ensemble member, as:

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\mathbf{x}_i^f), \quad (3)$$

in which the index  $i$  denotes the ensemble member index, with  $i=1, 2, \dots, N$ . Burgers *et al.* (1998) showed that the posterior covariance from this ensemble is systematically too small, even when the ensemble size goes to infinity. Indeed, when we calculate the posterior ensemble mean for an ensemble of  $N$  members as

$$\begin{aligned} \overline{\mathbf{x}^a} &= \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i^a = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i^f + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\overline{\mathbf{x}^f}) \\ &= (\mathbf{I} - \mathbf{K}\mathbf{H})\overline{\mathbf{x}^f} + \mathbf{K}\mathbf{y}^o, \end{aligned} \quad (4)$$

use

$$\begin{aligned} \mathbf{x}_i^a - \overline{\mathbf{x}^a} &= \mathbf{x}_i^f + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\mathbf{x}_i^f) - (\mathbf{I} - \mathbf{K}\mathbf{H})\overline{\mathbf{x}^f} - \mathbf{K}\mathbf{y}^o \\ &= (\mathbf{I} - \mathbf{K}\mathbf{H})(\mathbf{x}_i^f - \overline{\mathbf{x}^f}), \end{aligned} \quad (5)$$

and calculate the posterior sample covariance matrix

$$\begin{aligned} \mathbf{P}_e &= \frac{1}{N-1} \sum_{i=1}^N (\mathbf{x}_i^a - \overline{\mathbf{x}^a})(\mathbf{x}_i^a - \overline{\mathbf{x}^a})^T \\ &= (\mathbf{I} - \mathbf{K}\mathbf{H}) \frac{1}{N-1} \sum_{i=1}^N (\mathbf{x}_i^f - \overline{\mathbf{x}^f})(\mathbf{x}_i^f - \overline{\mathbf{x}^f})^T (\mathbf{I} - \mathbf{K}\mathbf{H})^T \\ &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}_e(\mathbf{I} - \mathbf{K}\mathbf{H})^T, \end{aligned} \quad (6)$$

in which  $\mathbf{B}_e$  is the prior sample covariance matrix. Comparing this to the Joseph form in Equation (2), we immediately see that, even in the limit of an infinite number of ensemble members, a term  $\mathbf{K}\mathbf{R}\mathbf{K}^T$  is missing.

Burgers *et al.* (1998) argue that the missing term is due to the fact that each ensemble member is updated with the same observation, leading them to suggest a perturbed observation scheme. Perturbing observations in an EnKF scheme was first advocated by Houtekamer and Mitchell (1998), using a Monte-Carlo argument and also found independently by Bennett *et al.* (1998) in their representer method, an observation-space variational scheme. The idea of perturbing observations goes back to Daley and Mayer (1986), who used it in observation system simulation experiments, and their use was extended by Houtekamer and Derome (1995) to generate forecast ensembles, and by Kitanidis (1995), who used it in a variational data-assimilation approach. It later found its way in 4D-Var for weather prediction by Žagar *et al.* (2005) and is now used operationally in the “Ensemble of Data Assimilations” scheme of the ECMWF, e.g., Isaksen *et al.* (2011).

In the perturbed-observations scheme, each ensemble member is updated with a perturbed observation

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K}(\mathbf{y}^o + \epsilon_i - \mathbf{H}\mathbf{x}_i^f), \quad (7)$$

in which  $\epsilon_i \sim N(0, \mathbf{R})$ . Since  $\epsilon_i$  and the prior ensemble members are independent draws from independent distributions, this would indeed lead to the correct covariance update in the limit of an infinite ensemble size:

$$\begin{aligned} \mathbf{x}_i^a - \overline{\mathbf{x}^a} &= \mathbf{x}_i^f + \mathbf{K}(\mathbf{y}^o + \epsilon_i - \mathbf{H}\mathbf{x}_i^f) - (\mathbf{I} - \mathbf{K}\mathbf{H})\overline{\mathbf{x}^f} - \mathbf{K}\mathbf{y}^o \\ &= (\mathbf{I} - \mathbf{K}\mathbf{H})(\mathbf{x}_i^f - \overline{\mathbf{x}^f}) + \mathbf{K}\epsilon_i, \end{aligned} \quad (8)$$

leading to

$$\begin{aligned} \mathbf{P}_e &= \frac{1}{N-1} \sum_{i=1}^N (\mathbf{x}_i^a - \overline{\mathbf{x}^a})(\mathbf{x}_i^a - \overline{\mathbf{x}^a})^T \\ &= (\mathbf{I} - \mathbf{K}\mathbf{H}) \frac{1}{N-1} \sum_{i=1}^N (\mathbf{x}_i^f - \overline{\mathbf{x}^f})(\mathbf{x}_i^f - \overline{\mathbf{x}^f})^T (\mathbf{I} - \mathbf{K}\mathbf{H})^T \\ &\quad + (\mathbf{I} - \mathbf{K}\mathbf{H}) \frac{1}{N-1} \sum_{i=1}^N (\mathbf{x}_i^f - \overline{\mathbf{x}^f})\epsilon_i^T \mathbf{K}^T + (\dots)^T \end{aligned}$$

$$+ \mathbf{K} \epsilon_i \epsilon_i^T \mathbf{K}^T \\ \approx (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}_e(\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{R}_e\mathbf{K}^T. \quad (9)$$

This expression converges to the Joseph form for infinite ensemble size. Note that one has to take care in how one takes the limit towards infinite ensemble size, and details can be found in e.g., Mandel *et al.* (2011)

The alternative methodology that leads to the same posterior mean and covariance perturbs the modelled observations, via

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K} \{ \mathbf{y}^o - (\mathbf{H}\mathbf{x}_i^f + \epsilon_i) \}, \quad (10)$$

in which  $\epsilon_i \sim N(0, \mathbf{R})$ . Repeating the analysis for the perturbed observation scheme given above with the sign of  $\epsilon$  changed does show that this scheme leads to the same posterior statistics. This version of the stochastic EnKF has been used more and more (e.g., the review on ensemble methods in Vetra-Carvalho *et al.*, 2018). The first to suggest this scheme was Hodyss (2011), as far as this author is aware. His idea originated from studying non-Gaussian extensions of the EnKF. However, he sees no preference for either scheme for the Gaussian case. The paper by Nott *et al.* (2012) is slightly confusing, and seems to mix ideas from both schemes. Snyder (2015) presented a clear preference for the perturbed modelled observations scheme based on his analysis of the EnKF approximation of the Best Linear Unbiased Estimator (BLUE).

Is there any preference for either of these two schemes? One could argue no, because both lead to the same answer. However, as with any analysis method, it would be good to understand where the perturbations come from. In Burgers *et al.* (1998), the reason for perturbing observations was after the fact: doing that leads to the correct result, and no principled reason was provided. Houtekamer and Mitchell (1998) suggest that observation perturbations are needed because the EnKF is a Monte-Carlo method, while Snyder (2015) argues that the BLUE equations directly suggest that one should perturb modelled observations, and that the BLUE in non-Gaussian situations provides further evidence for this choice.

The contribution of this paper is that all existing arguments for either scheme will be critically discussed, new arguments from ensemble verification methods, from Bayes' Theorem, and from a reanalysis of the EnKF as a BLUE are provided, and a new principled derivation of the stochastic EnKF will be presented. Furthermore, a numerical example is included to demonstrate the influence of the choice of perturbation of the posterior when the likelihood is skewed. The conclusion is that preference should be given to perturbing modelled observations as strong arguments exist for that scheme, while no existing arguments remain and no new arguments

have been found in favour of the scheme that perturbs observations.

In the next section arguments in favour of the perturbed observation scheme are presented and discussed. Then arguments in favour of perturbing modelled observations are provided, and a new consistent derivation of the stochastic EnKF is given. This is followed by a numerical non-Gaussian example that also supports perturbing modelled observations. A summarizing and concluding section closes the paper. To keep the paper focussed, we will not discuss issues like inbreeding, localisation and inflation, or efficient updating schemes as they are not directly relevant for the present discussion.

## 2 | ARGUMENTS IN FAVOUR OF PERTURBING OBSERVATIONS

Houtekamer and Mitchell (1998), Burgers *et al.* (1998), and Bennett *et al.* (1998) introduce observation perturbations by arguing implicitly or explicitly that it is needed for a full Monte-Carlo description of the posterior. However, there is an enormous volume of literature on Markov-Chain Monte-Carlo methods, such as Gibbs Sampling, Metropolis Hastings Sampling, Langevin Sampling and Particle Filters, and none of these methods need to perturb observations, e.g., Doucet *et al.* (2000), Robert and Casella (2004), van Leeuwen (2009), Bocquet *et al.* (2010), van Leeuwen *et al.* (2015), van Leeuwen *et al.* (2019). It is known that all these methods are providing the correct posterior pdf in the limit of an infinitely large ensemble size, showing that this argument does not hold.

It is interesting that no derivation of a perturbed observation scheme exists. Papers that “derive” the scheme only show that using the update equation with perturbed observations leads to the correct posterior covariance; none of them provides a principled derivation for this scheme (e.g., the papers mentioned in the introduction and Mandel *et al.*, 2011). This author has also not been able to provide such a derivation.

## 3 | ARGUMENTS IN FAVOUR OF PERTURBING MODELLED OBSERVATIONS

In the following we provide three reasons why perturbing modelled observations is preferable over perturbing observations.

### 3.1 | Reasons from ensemble verification

In ensemble verification methods, as in data assimilation, we compare observations with the ensemble member

equivalents of those observations. Each prior ensemble member is interpreted as a random draw from the prior pdf  $p(\mathbf{x})$  (e.g., the rank histogram literature, such as Hamill, 2001). Also the true state of the system  $\mathbf{x}_{\text{true}}$  is assumed to be a random draw from that pdf. Hence, each ensemble member is statistically indistinguishable from the true state.

The observations  $\mathbf{y}^o$  arise from applying the observation operator to the true state of the system  $\mathbf{H}\mathbf{x}_{\text{true}}$ , followed by a perturbation with a random measurement error  $\epsilon \sim N(0, \mathbf{R})$ . If each ensemble member is statistically indistinguishable from the true state, we should treat the measurement of that member the same as a measurement of the true state, so the ensemble member equivalent of the observation is  $\mathbf{H}\mathbf{x}_i + \epsilon_i$  in which  $\epsilon_i \sim N(0, \mathbf{R})$ . Note that the equivalent of the measurement of an ensemble member to a real observation is not  $\mathbf{H}\mathbf{x}_i$ : adding  $\epsilon_i$  is essential. The consequence of this reasoning in favour of perturbing modelled observations has not been highlighted earlier.

The argument can also be reformulated against perturbing observations. The observations are obtained by adding measurement noise to applying the observation operator to the true state of the system. Nature does this random draw for us. Hence the observation is derived from adding a perturbation to the true system. Why further perturb a quantity that is itself a perturbation from the true system? Also, note that any observation, *after the observation has been made*, is a biased estimate of the true system because of the non-zero measurement error. Why would one want to create an ensemble around a biased variable? We come back to this reasoning when discussing arguments from the BLUE below.

### 3.2 | Reasons from Bayes' Theorem

Bayes' Theorem reads:

$$p(\mathbf{x}|\mathbf{y}^o) = \frac{p(\mathbf{y}^o|\mathbf{x})}{p(\mathbf{y}^o)} p(\mathbf{x}), \quad (11)$$

which expresses the pdf of the state of the system *given a set of observations*  $\mathbf{y}^o$ , in terms of the likelihood and the prior. The likelihood  $p(\mathbf{y}^o|\mathbf{x})$  is not the pdf of the observations, as these are given. It is an unnormalized function of the state  $\mathbf{x}$  given a set of observations. Similarly,  $p(\mathbf{y}^o)$  is a constant, as, again,  $\mathbf{y}^o$  is a given vector. Bayes' Theorem is a point-wise update for each possible value of the state vector, and the full solution to the data assimilation problem is the full posterior pdf of the state  $p(\mathbf{x}|\mathbf{y}^o)$ . (e.g. van Leeuwen 2015 gives a thorough discussion of the meaning of the likelihood in Bayes' Theorem.) In Bayes' Theorem the state  $\mathbf{x}$  and the observations  $\mathbf{y}^o$  are not treated equally: one is a

random variable, and the other is fixed. This is why Bayes' Theorem goes beyond the simple notion of a conditional pdf. It is derived from the identity  $p(\mathbf{x}|\mathbf{y})p(\mathbf{y}) = p(\mathbf{y}|\mathbf{x})p(\mathbf{x})$  from probability theory, but then it is used as an update equation for the pdf of the state when observations  $\mathbf{y}$  are given as  $\mathbf{y}^o$ , and enters the field of data assimilation, and Bayesian Inference in general.

How can this be compatible with a perturbed observation scheme to find the correct posterior covariance? Burgers *et al.* (1998) explore the expression for the posterior covariance as an expectation over the squared average deviation of the state from its mean. Using Bayes' Theorem, the posterior covariance is the squared average deviation of the state from its mean, where the integral is weighted with the posterior pdf:

$$\mathbf{P} = E_{\mathbf{x}|\mathbf{y}^o} [(\mathbf{x} - \mathbf{x}^a)(\mathbf{x} - \mathbf{x}^a)^T], \quad (12)$$

in which  $E_{\mathbf{x}|\mathbf{y}^o}[\dots]$  denotes the expectation with respect to the pdf of  $\mathbf{x}|\mathbf{y}^o$ . Burgers *et al.* (1998) use an 'overbar' notation to denote the expectation over the deviations from its mean, but it is unclear what the underlying pdf is. The only consistent interpretation is that they use the joint pdf  $p(\mathbf{x}, \mathbf{y})$ , and hence their derivation is based on the BLUE, not on Bayes' Theorem, as shown in the next section.

Appendix B shows that it is possible to derive the correct posterior covariance

$$\mathbf{P} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}(\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{R}\mathbf{K}^T \quad (13)$$

using the posterior pdf in the expectation, so using Bayes' Theorem. This derivation is new, and this is the first time that a stochastic EnKF is shown to be consistent with Bayes' Theorem. The important point of this derivation is that observations are treated as given constants, not as random variables.

### 3.3 | Reasons from the BLUE

Snyder (2015) showed that a stochastic EnKF can be derived using the Best Linear Unbiased Estimator formalism. This differs from Bayes' Theorem in that one searches for an analysis state that is an unbiased linear combination of prior state and observations with minimal trace of the analysis covariance. The BLUE imposes that this trace is minimal irrespective of the actual value of the observations, or, in other words, is minimal averaged over all possible realizations of the actual observation. Hence in this section  $\mathbf{y}$  is a random variable, and the expectation in the expression for the posterior covariance uses the joint pdf  $p(\mathbf{x}, \mathbf{y})$  of state and observations, and not the posterior pdf  $p(\mathbf{x}|\mathbf{y})$  as in Bayes' Theorem.



Snyder (2015) shows that the BLUE estimate is equal to

$$\mathbf{x}_{\text{BLUE}} = \mathbf{x}^f + \mathbf{K}(\mathbf{y}^o - E[\mathbf{y}]), \quad (14)$$

in which  $\mathbf{y}^o$  is the actual observation and the expectation  $E[\mathbf{y}]$  is taken with respect to  $p(\mathbf{x}, \mathbf{y})$ , or  $p(\mathbf{y})$  in this case as realizations of  $\mathbf{y}$  are independent from realizations of  $\mathbf{x}$ . It is important to realize that  $E[\mathbf{y}]$  is not the modelled observation. It is a quantity defined by  $E[\mathbf{y}] = \int \mathbf{y} p(\mathbf{x}, \mathbf{y}) d\mathbf{x}$ .

The following makes the wording in Snyder (2015) explicit. The expectation  $E[\mathbf{y}]$  is defined via

$$E[\mathbf{y}] = \int \mathbf{y} p(\mathbf{x}, \mathbf{y}) d\mathbf{x} = \int \mathbf{y} p(\mathbf{y}|\mathbf{x}) p(\mathbf{x}) d\mathbf{x}.$$

If we now use  $p(\mathbf{y}|\mathbf{x}) = N(\mathbf{H}\mathbf{x}, \mathbf{R})$  and  $p(\mathbf{x}) = N(\mathbf{x}^f, \mathbf{B})$ , the integral can be evaluated as  $E[\mathbf{y}] = \mathbf{H}\mathbf{x}^f$ , consistent with the Kalman Filter assumptions. The connection with a stochastic EnKF can be made by using an ensemble representation with  $N$  members for  $p(\mathbf{x})$  in the integral, leading to

$$E[\mathbf{y}] = \int \mathbf{y} p(\mathbf{y}|\mathbf{x}) p(\mathbf{x}) d\mathbf{x} \approx \frac{1}{N} \sum_{i=1}^N \int \mathbf{y} p(\mathbf{y}|\mathbf{x}_i) d\mathbf{y}, \quad (15)$$

with the  $\mathbf{x}_i$  being draws from the prior. If we now draw one  $\mathbf{y}_i$  for each  $\mathbf{x}_i$ , so we draw  $\mathbf{y}_i$  from  $p(\mathbf{y}|\mathbf{x}_i) = N(\mathbf{H}\mathbf{x}_i, \mathbf{R})$ , we find

$$E[\mathbf{y}] \approx \frac{1}{N} \sum_{i=1}^N \mathbf{y}_i = \frac{1}{N} \sum_{i=1}^N \mathbf{H}\mathbf{x}_i + \epsilon_i. \quad (16)$$

This clearly demonstrates the need for perturbing modelled observations.

We now present an alternative derivation following the steps taken in the derivation using Bayes' Theorem that will allow for more insight in the role of the observations, and allows for a direct connection with Burgers *et al.* (1998). In this case we define, as implicitly done in Burgers *et al.* (1998),

$$\begin{aligned} \mathbf{P} &= \int (\mathbf{x} - \mathbf{x}^a)(\mathbf{x} - \mathbf{x}^a)^T p(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y} \\ &= E_{\mathbf{x}, \mathbf{y}} [(\mathbf{x} - \mathbf{x}^a)(\mathbf{x} - \mathbf{x}^a)^T] \end{aligned} \quad (17)$$

and Appendix C shows how one derives the Joseph form of the posterior covariance from this expression. We now trace back in that derivation whether the observations themselves, or the modelled observations from the ensemble, would be perturbed in a stochastic EnKF. The observational part appears as an integral of the form

(Equation (C2))

$$\int \left[ \int \{\mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x})\} \{\mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x})\}^T p(\mathbf{y}|\mathbf{x}) d\mathbf{y} \right] p(\mathbf{x}) d\mathbf{x}. \quad (18)$$

We now implement a Monte-Carlo sampling method for this integral, as in a stochastic version of the EnKF. We first draw  $N$  samples  $\mathbf{x}_i$  from  $p(\mathbf{x})$ , leading to:

$$\begin{aligned} &\int \left[ \int \{\mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x})\} \{\mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x})\}^T p(\mathbf{y}|\mathbf{x}) d\mathbf{y} \right] p(\mathbf{x}) d\mathbf{x} \\ &\approx \frac{1}{N} \sum_{i=1}^N \int \{\mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}_i)\} \{\mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}_i)\}^T p(\mathbf{y}|\mathbf{x}_i) d\mathbf{y}. \end{aligned} \quad (19)$$

To evaluate this further we draw a sample from  $p(\mathbf{y}|\mathbf{x}_i)$  for each  $i$ . This is a pdf with mean  $\mathbf{H}\mathbf{x}_i$  and covariance  $\mathbf{R}$ , and the important observation is that this pdf is not centred at the actual observations. The result of such a draw is  $\mathbf{y}_i = \mathbf{H}\mathbf{x}_i + \epsilon_i$ , in which  $\epsilon_i$  is drawn from  $N(0, \mathbf{R})$ . Hence we perturb  $\mathbf{H}\mathbf{x}_i$  and not the actual observation  $\mathbf{y}^o$ , consistent with Snyder (2015).

(Nott *et al.* (2012) use a similar reasoning, but their wording is slightly confusing. They introduce the stochastic EnKF in two steps. First they draw samples  $(\mathbf{x}_i, \mathbf{y}_i)$  from  $p(\mathbf{y}|\mathbf{x})p(\mathbf{x})$  by first drawing from  $p(\mathbf{x})$  followed by a draw from each  $p(\mathbf{y}|\mathbf{x}_i)$ . This is the same as done here. Then they mention that “*The next important idea in the ensemble Kalman Filter is to adjust the forecast ensemble ... There are several ways to do this, but one common way that provides a connection with ABC methods is the following perturbed observations scheme*”, suggesting that the connection between the draws  $\mathbf{y}_i$  and perturbations in the innovation is actually not made.)

We can analyze the BLUE derivation also by interchanging the integrations in (18), and first sample from  $p(\mathbf{y})$ :

$$\begin{aligned} &\int \left[ \int \{\mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x})\} \{\mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x})\}^T p(\mathbf{x}|\mathbf{y}) d\mathbf{x} \right] p(\mathbf{y}) d\mathbf{y} \\ &\approx \frac{1}{N} \sum_{i=1}^N \int \{\mathbf{K}(\mathbf{y}_i - \mathbf{H}\mathbf{x})\} \{\mathbf{K}(\mathbf{y}_i - \mathbf{H}\mathbf{x})\}^T p(\mathbf{x}|\mathbf{y}_i) d\mathbf{x}. \end{aligned} \quad (20)$$

Does this lead to a perturbed observation scheme? The answer is again no.  $p(\mathbf{y})$  is a pdf with mean  $\mathbf{H}\mathbf{x}^f$  and covariance  $\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}$ , as can be seen directly from the identity  $p(\mathbf{y}) = \int p(\mathbf{y}|\mathbf{x})p(\mathbf{x}) d\mathbf{x}$ . Hence this pdf  $p(\mathbf{y})$  is not centred at the actual observations. In fact, we can write  $\mathbf{y}_i = \mathbf{H}\mathbf{x}^f + \xi_i$  in which  $\xi_i \sim N(0, \mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})$ . To evaluate the integral further, we need to draw samples from  $p(\mathbf{x}|\mathbf{y}_i)$ .

This looks very much like an ensemble method in which each ensemble member sees a different observation realization, as in a perturbed observation scheme. This is true, with the important difference that the  $\mathbf{y}_i$  are centred on  $\mathbf{H}\mathbf{x}^f$  and not on the actual observations.

### 3.4 | Discussion

The above presented three reasons to perturb the modelled observations and not the observations themselves. No reasons could be found by this author to perturb the observations, and this paper partly serves to correct his mistake in Burgers *et al.* (1998).

All three reasons presented above in favour of perturbing modelled observations hold equally well for linear ensemble smoothers as these are just state-space extensions to the time domain of the Ensemble Kalman Filter. It is expected that perturbing modelled observations instead of perturbing observations is also beneficial for non-Gaussian situations where iterative variants of the update scheme are used, such as stochastic iterative ensemble filters and smoothers and the “Ensemble of Data Assimilations” scheme (e.g., Žagar *et al.*, 2005 and Isakson *et al.*, 2011) as these can be seen as Gauss–Newton iterations of the linear scheme discussed here.

## 4 | A CONSISTENT DERIVATION OF A STOCHASTIC ENKF

The basic idea behind ensemble or particle methods is that one tries to draw samples from the posterior pdf, starting from samples from the prior pdf. Several methods are available, and a short overview is given in Appendix A to understand better where Ensemble Kalman Filters fit in, which is needed for our discussion.

When the prior and posterior are Gaussian, a one-step transformation (method 4 in Appendix A) can be made exact. Deterministic Ensemble Kalman Filters explore this by using an ensemble square root of the prior and posterior covariance. For instance, the Ensemble Transform Kalman Filter (Bishop *et al.*, 2001) transforms the so-called ensemble perturbation matrices via  $\mathbf{X}^a = \mathbf{X}\mathbf{T}$  in which  $\mathbf{T}$  is a transformation matrix. The ensemble mean is transformed similarly, and indeed the ensemble is transformed from the prior to the posterior in one deterministic step.

The stochastic version of the Ensemble Kalman Filter is special because it draws samples directly from the Gaussian posterior pdf, and is an example of scheme 6 in Appendix A. If the mean and covariance of the posterior are given by  $\mathbf{x}^a$  and  $\mathbf{P}$ , respectively, we can generate

samples via

$$\mathbf{x}_i^a = \mathbf{x}^a + \mathbf{P}^{1/2} \boldsymbol{\xi}_i, \quad (21)$$

in which  $\boldsymbol{\xi}_i \sim N(0, \mathbf{I})$ . Because these samples are direct independent samples from the posterior, there is no need for any weighting, or no need for extra moves of the ensemble members. (Not realising this has led to some confusion in the literature, e.g., Morzfeld *et al.*, 2017 who discuss using the EnKF as proposal in a particle filter scheme, method 2 in Appendix A. They did not mention that the stochastic version of the EnKF is developed as a method to directly sample from the Gaussian posterior.)

To use the sampling scheme (Equation (21)) we need to determine  $\mathbf{P}^{1/2}$ . Since  $\mathbf{P}$  is a covariance matrix, it is symmetric positive semi-definite and it has one unique symmetric positive semi-definite square root. In the stochastic version of the EnKF, one does not use this symmetric square root because there is no analytical expression for the posterior covariance as one symmetric matrix. Instead, a non-symmetric square root can be used. Starting from the Joseph form of the posterior covariance,

$$\mathbf{P} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}(\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{R}\mathbf{K}^T, \quad (22)$$

we note that the posterior covariance matrix is a summation of two symmetric matrices. Using this Joseph form for  $\mathbf{P}$ , the natural square-root matrix is given by

$$\mathbf{P}^{1/2} = \{(\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}^{1/2}, -\mathbf{K}\mathbf{R}^{1/2}\}, \quad (23)$$

where the minus sign in front of the second block of the covariance follows from the minus sign in front of  $\mathbf{K}$  in the factor  $\mathbf{x} - \mathbf{x}^a = \mathbf{x} - \mathbf{x}^f - \mathbf{K}(\mathbf{y}^o - \mathbf{H}\mathbf{x}^f)$  in the definition of  $\mathbf{P}$ . It can be traced back to the minus sign in front of  $\epsilon$  in Equation (C4). Also Appendix B shows in Equations (B1) and (B4) that the natural factor appearing in the equations is  $\mathbf{H}\mathbf{x} - \mathbf{y}^o$ , which corresponds to a negative sign for  $\epsilon$ . It should be noted that, although the minus sign in Equation (23) is more straightforward and perhaps more natural, this derivation does not explicitly force us to use the minus sign. All equations in this section are equally valid up to this point with the positive sign in front of  $\mathbf{K}\mathbf{R}^{1/2}$ . The formal justification for the minus sign comes from the arguments laid down in the previous section and in the next section.

Matrix  $\mathbf{P}^{1/2}$  is of size  $N_x \times (N_x + N_y)$ , with  $N_x$  the dimension of the state vector and  $N_y$  the dimension of the observation vector, and hence the random vector  $\boldsymbol{\xi}_i$  is a vector in  $\mathcal{R}^{N_x + N_y}$ . We can split this vector into two parts  $\boldsymbol{\xi}_i = (\boldsymbol{\xi}_i^x, \boldsymbol{\xi}_i^y)^T$ , where the superscript refers to the space each part belongs to.



Using this expression for the square root of  $\mathbf{P}$ , we find for the posterior samples:

$$\begin{aligned}\mathbf{x}_i^a &= \mathbf{x}^a + \mathbf{P}^{1/2} \boldsymbol{\xi}_i \\ &= (\mathbf{I} - \mathbf{KH})\mathbf{x}_f + (\mathbf{I} - \mathbf{KH})\mathbf{B}^{1/2} \boldsymbol{\xi}_i^x + \mathbf{K}\mathbf{y}^o - \mathbf{KR}^{1/2} \boldsymbol{\xi}_i^y.\end{aligned}\quad (24)$$

If we now realise that we do have samples from the prior as  $\mathbf{x}_i^f = \mathbf{x}^f + \mathbf{B}^{1/2} \boldsymbol{\xi}_i^x$  and we define  $\boldsymbol{\epsilon}_i = \mathbf{R}^{1/2} \boldsymbol{\xi}_i^y$ , we can write

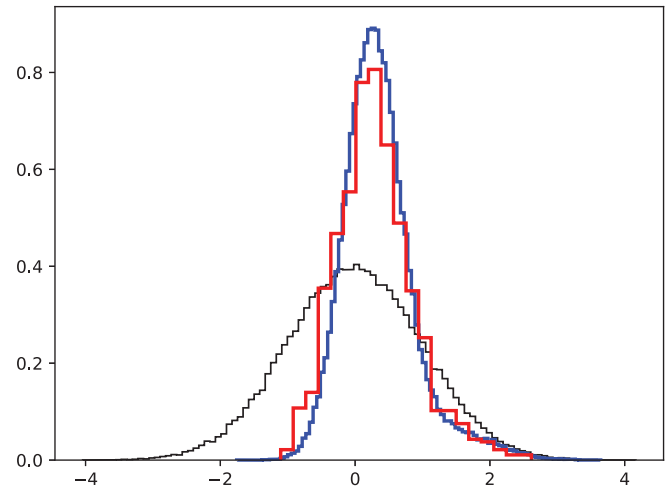
$$\begin{aligned}\mathbf{x}_i^a &= (\mathbf{I} - \mathbf{KH})\mathbf{x}_i^f + \mathbf{K}\mathbf{y}^o - \mathbf{K}\boldsymbol{\epsilon}_i \\ &= \mathbf{x}_i^f + \mathbf{K} \{ \mathbf{y}^o - (\mathbf{H}\mathbf{x}_i^f + \boldsymbol{\epsilon}_i) \}.\end{aligned}\quad (25)$$

The above derivation is perhaps the most simple derivation of the stochastic version of the EnKF produced so far, based on its interpretation as a scheme that directly samples from the posterior pdf. In principle, one could subtract the  $\boldsymbol{\epsilon}_i$  from the observations instead of adding them to the measured model states, but the ensemble-verification argument in section 3.2, and also the other arguments, are against this choice.

## 5 | WEAKLY NON-GAUSSIAN STOCHASTIC ENKF

Direct practical differences between perturbing observations or perturbing modelled observations will become apparent when the observation errors are not Gaussian or the observation operator is nonlinear. As mentioned in Hodyss (2011) and Snyder (2015), a skewed pdf of the observation errors  $\boldsymbol{\epsilon}$  will lead to a skewed posterior ensemble. The stochastic version of the EnKF formulation in which we perturb the  $\mathbf{H}\mathbf{x}_i$  leads to a skewness of the correct sign, while the “perturbed observation” approach does not. One could argue that an EnKF should only be used in a Gaussian setting, but real-world problems are never Gaussian. Instead, one could argue that a scheme that behaves well both in the Gaussian and in the weakly non-Gaussian regime is preferable.

To understand why the perturbed modelled observation scheme is preferable, assume that the prior is symmetric and the likelihood is skewed. Specifically, let us assume that the observation errors are negatively skewed, meaning that negative values of  $\boldsymbol{\epsilon}$  tend to be of larger magnitude than positive ones. To understand the argument it is important to follow the sign of epsilon through the analysis scheme. In the likelihood we need to evaluate  $\boldsymbol{\epsilon} = \mathbf{y}^o - \mathbf{x}$ , in which  $\mathbf{y}^o$  is given. Hence the likelihood as function of  $\mathbf{x}$  will be positively skewed. This means that the posterior will also be positively skewed. Let us now use the stochastic



**FIGURE 1** Numerical illustration of an EnKF on a skewed likelihood for the perturbed modelled observation version. The prior is depicted as the broad Gaussian with the black line, the true posterior as the blue line, and the stochastic EnKF posterior as the red line. Note that this stochastic EnKF and the posterior have the same positive skewness

EnKF formulation as advocated in the previous section:

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K} \{ \mathbf{y}^o - (\mathbf{H}\mathbf{x}_i^f + \boldsymbol{\epsilon}_i) \}.\quad (26)$$

Since the  $\boldsymbol{\epsilon}$  will be negatively skewed, the posterior samples will be positively skewed, as they should be when they are to be samples of the posterior pdf. In contrast, the “perturbed observations” approach will lead to a posterior sample skewness of the wrong sign.

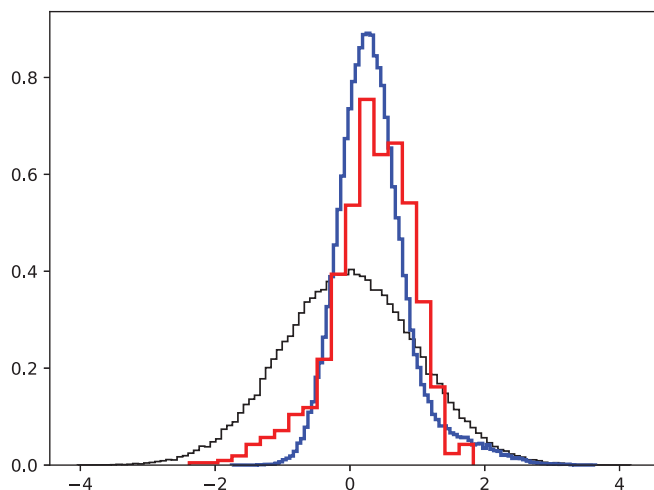
An illustrative numerical example is discussed next. Assume a zero-dimensional system with a Gaussian prior  $N(0,1)$ . An observation with value  $y$  is obtained, with error  $\boldsymbol{\epsilon} \sim p(\boldsymbol{\epsilon})$  in which

$$p(\boldsymbol{\epsilon}) = \alpha_1 N(\boldsymbol{\epsilon}_1, \sigma_1^2) + (1 - \alpha_1) N(\boldsymbol{\epsilon}_2, \sigma_2^2).\quad (27)$$

We ensure that  $E_{\boldsymbol{\epsilon}}[\boldsymbol{\epsilon}] = \alpha_1 \boldsymbol{\epsilon}_1 + \alpha_2 \boldsymbol{\epsilon}_2 = 0$  by choosing  $\boldsymbol{\epsilon}_2 = -\alpha_1 \boldsymbol{\epsilon}_1 / (1 - \alpha_1)$ . We choose  $\alpha_1 > 0.5$ , such that  $p(\boldsymbol{\epsilon})$  is negatively skewed, meaning that negative values of  $\boldsymbol{\epsilon}$  tend to be of larger magnitude than positive ones.

For this specific example we choose  $x^f = 0$ ,  $\sigma_f^2 = 1$ ,  $\alpha_1 = 0.9$ ,  $\boldsymbol{\epsilon}_1 = 0.2$ ,  $\sigma_1^2 = 0.2$ ,  $\sigma_2^2 = 0.7$ , and  $y^o = 0.5$ .

To generate the true posterior, we draw 100,000 samples from the posterior density, which, since the prior is Gaussian and the likelihood is a Gaussian mixture, is also a Gaussian mixture. We run a stochastic EnKF with perturbed modelled observations by drawing 1,000 samples from the Gaussian prior. The observation error covariance is determined from the 100,000 samples from  $p(\boldsymbol{\epsilon})$ .



**FIGURE 2** Numerical illustration of an EnKF on a skewed likelihood for the perturbed observation version. The prior is depicted as the broad Gaussian with the black line, the true posterior as the blue line, and the stochastic EnKF posterior as the red line. The same random realizations for the perturbations are used as in Figure 1. Note that this stochastic EnKF has a skewness of the wrong sign compared to the posterior

Figure 1 contains the result of the numerical experiment, showing the broad Gaussian prior, the true posterior, and the stochastic EnKF posterior. The positive skewness of the posterior and the stochastic EnKF are directly visible.

Figure 2 shows the same picture but using the 'perturbed observation' EnKF, as in Equation (7), clearly showing that the EnKF posterior has a skewness of the wrong sign. This negative result should not come as a surprise because, as argued above, the method is inconsistent with Bayes' Theorem which describes this situation exactly.

## 6 | SUMMARY AND CONCLUSIONS

Both the perturbed-observation EnKF and the perturbed modelled observation EnKF give rise to a statistically correct posterior mean and covariance when the prior and the likelihood are Gaussian in the state of the system. However, arguments from ensemble verification methods, Bayes' Theorem and the BLUE all suggest that one should perturb the modelled observations. No such arguments have been found for the perturbed observation scheme, and the Monte-Carlo reason proposed earlier has been invalidated. Furthermore, it was shown explicitly that a skewed likelihood also has a clear preference for the perturbed modelled observation scheme.

This paper thus concludes that the correct interpretation of a stochastic EnKF is one in which one perturbs the modelled observations and not the actual observations.

A simple derivation of the stochastic EnKF as a scheme that draws samples directly from the posterior and perturbs modelled observations is provided.

The results of this paper carry over without reservation to linear ensemble smoothers as these are just state-space extensions to the time domain of the Ensemble Kalman Filter. Iterative variants of those, such as stochastic iterative ensemble smoothers (e.g., Gu and Oliver, 2007; Emerick and Reynolds, 2013), and the "Ensemble of Data Assimilations" scheme advocated by operational centres (e.g., Žagar *et al.*, 2005; Isaksen *et al.*, 2011), are also expected to benefit from a consistent treatment of the linear problem as advocated in this paper since the same arguments as displayed here can be used on these schemes.

## ACKNOWLEDGEMENTS

The author thanks the European Research Council for funding via the CUNDA project under number 694509.

## ORCID

Peter Jan van Leeuwen  <https://orcid.org/0000-0003-2325-5340>

## REFERENCES

- Ades, M. and van Leeuwen, P.J. (2013) An exploration of the equivalent weights particle filter. *Quarterly Journal of the Royal Meteorological Society*, 139, 820–840.
- Anderson, J.L. (2001) An ensemble adjustment Kalman filter for data assimilation. *Monthly Weather Review*, 129, 2884–2903.
- Asch, M., Bocquet, M. and Nodet, M. (2016) *Data assimilation, methods, algorithms, and applications*. SIAM, Philadelphia, PA.
- Bennett, P.F., Chua, B.S., Harrison, D.E. and McPhaden, M.J. (1998) Generalized inversion of tropical atmosphere–ocean (tao) data and a coupled model of the tropical Pacific. *Journal of Climate*, 11, 1768–1792.
- Bishop, C.H., Etherton, B.J. and Majumdar, S.J. (2001) Adaptive sampling with the ensemble transform Kalman Filter. Part I: theoretical aspects. *Monthly Weather Review*, 129, 420–436.
- Bocquet, M., Pires, C.A. and Wu, L. (2010) Beyond Gaussian statistical modeling in geophysical data assimilation. *Monthly Weather Review*, 138. <https://doi.org/10.1175/2010MWR3164.1>.
- Burgers, G., van Leeuwen, P.J. and Evensen, G. (1998) Analysis scheme in the ensemble Kalman filter. *Monthly Weather Review*, 126(6), 1719–1724.
- Daley, R. and Mayer, T. (1986) Estimates of global analysis error from the global weather experiment observational network. *Monthly Weather Review*, 114, 1642–1653.
- Doucet, A., Godsill, S. and Andrieu, C. (2000) On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10, 197–208.
- Emerick, A. and Reynolds, A. (2013) Ensemble smoother with multiple data assimilations. *Computers and Geosciences*, 55, 3–15.
- Evensen, G. (1994) Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99, 10143–10162.

- Gu, Y. and Oliver, D. (2007) An iterative ensemble kalman filter for multiphase fluid flow data assimilation. *SPE Journal*, 12, 438–446. <https://doi.org/10.2118/108438-PA>.
- Hamill, T.M. (2001) Interpretation of rank histograms for verifying ensemble forecasts. *Monthly Weather Review*, 129, 550–560.
- Hodyss, D. (2011) Ensemble state estimation for nonlinear systems using polynomial expansions in the innovation. *Monthly Weather Review*, 139. <https://doi.org/10.1175/2011MWR3558.1>.
- Houtekamer, P.L. and Derome, J. (1995) Methods for ensemble prediction. *Monthly Weather Review*, 123, 2181–2196.
- Houtekamer, P.L. and Mitchell, H.L. (1998) Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review*, 126, 796–811.
- Isaksen, L., Bonavita, M., Buizza, R., Fisher, M., Haseler, J., Leutbecher, M. and Raynaud, L. (2011) *Ensemble of data assimilations at ECMWF*. Technical Memorandum 636, ECMWF, Reading, UK.
- Kitanidis, P. (1995) Quasi-linear geostatistical theory for inversion. *Water Resources Research*, 31, 2411–2419.
- Liu, Q. and Wang, D. (2016). Stein variational gradient descent: a general purpose Bayesian inference algorithm. NIPS 2016. arXiv:1608.04471.
- Mandel, J., Cobb, L. and Breezley, J. (2011) On the convergence of the ensemble Kalman filter. *Applied Mathematics*, 56, 533–541. <https://doi.org/10.1007/s10492-011-0031-2>.
- Morzfeld, M., Hodyss, D. and Snyder, C. (2017) What the collapse of the Ensemble Kalman Filter tells us about particle filters. *Tellus A*, 69. <https://doi.org/10.1080/16000870.2017.1283809>.
- Moselhy, T.E. and Marzouk, Y. (2012) Bayesian inference with optimal maps. *Journal of Computational Physics*, 231(23), 7815–7850. <https://doi.org/10.1016/j.jcp.2012.07.022>.
- Nott, D., Marshall, L. and Ngoc, T. (2012) The Ensemble Kalman Filter is an ABC algorithm. *Statistics and Computing*, 22, 1273–1276. <https://doi.org/10.1007/s11222-011-9300-x>.
- Pulido, M. and van Leeuwen, P.J. (2019) Sequential Monte Carlo with kernel embedded mappings: The mapping particle filter. *Journal of Computational Physics*, 396, 400–415. <https://doi.org/10.1016/j.jcp.2019.06.060>.
- Reich, S. (2011) A dynamical systems framework for intermittent data assimilation. *BIT Numerical Mathematics*, 51, 235–249.
- Robert, C.P. and Casella, G. (2004) *Monte Carlo Statistical Methods*. Springer, New York, NY. <https://doi.org/10.1007/978-1-4757-4145-2>
- Snyder, C. (2015) *Advanced Data Assimilation for Geosciences, Lecture Notes of the Les Houches School of Physics*, Vol. 9, pp. 75–120.
- van Leeuwen, P.J. (2009) Particle filtering in geophysical systems. *Monthly Weather Review*, 137, 4089–4114.
- van Leeuwen, P.J. (2015) Representation errors and retrievals in linear and nonlinear data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 141, 1612–1623. <https://doi.org/10.1002/qj.2464>.
- van Leeuwen, P.J. and Evensen, G. (1996) Data assimilation and inverse methods in terms of a probabilistic formulation. *Monthly Weather Review*, 124, 2898–2913.
- van Leeuwen, P.J., Cheng, Y. and Reich, S. (2015). Nonlinear data assimilation, in *Frontiers in Applied Dynamical Systems: Reviews and Tutorials 2*. Jones, C.K.R.T., Sandstede, B., Young, L.-S. (eds), Springer, Basel, Switzerland.
- van Leeuwen, P.J., Kunsch, H.R., Nerger, L., Potthast, R. and Reich, S. (2019) Particle filters for high-dimensional geoscience applications: A review. *Quarterly Journal of the Royal Meteorological Society*, 145, 2335–2365. <https://doi.org/10.1002/qj.3551>.
- Vetra-Carvalho, S., van Leeuwen, P.J., Nerger, L., Barth, A., Altaf, U., Brasseur, P., Kirchgeßner, P. and Beckers, J.M. (2018) State-of-the-art stochastic data assimilation methods for high-dimensional non-Gaussian problems. *Tellus A*. <https://doi.org/10.1080/16000870.2018.1445364>.
- Villani, C. (2008) *Optimal Transport: Old and New*. Springer, New York, NY.
- Whitaker, J.S. and Hamill, T.M. (2002) Ensemble data assimilation without perturbed observations. *Monthly Weather Review*, 130, 1913–1924.
- Žagar, N., Andersson, E. and Fisher, M. (2005) Balanced tropical data assimilation based on a study of equatorial waves in ECMWF short-range forecast errors. *Quarterly Journal of the Royal Meteorological Society*, 131, 987–1011. <https://doi.org/10.1256/qj.04.54>.
- Zhu, M., van Leeuwen, P.J. and Amezcua, J. (2016) Implicit equal-weights particle filter. *Quarterly Journal of the Royal Meteorological Society*, 142, 1904–1919. <https://doi.org/10.1002/qj.2784>.

**How to cite this article:** van Leeuwen P.J. A consistent interpretation of the stochastic version of the Ensemble Kalman Filter. *Q J R Meteorol Soc.* 2020;146:2815–2825. <https://doi.org/10.1002/qj.3819>

## APPENDICES

### A. Methods for generating posterior samples

To better understand the sampling methodology of a stochastic EnKf, a short overview of posterior sampling methods is given below. Sampling methods can:

1. Use the prior samples directly as samples from the posterior, which are then weighted by the likelihood, as in the standard particle filter (e.g., Doucet *et al.*, 2000, van Leeuwen, 2009, van Leeuwen *et al.*, 2015), or
2. Move the prior samples towards the high-likelihood regions of state space via a proposal step, leading to samples that are closer to the observations with weights that are the product of the new likelihood and a proposal weight (e.g., Doucet *et al.*, 2000, van Leeuwen, 2009, van Leeuwen *et al.*, 2015), or
3. Explore tempering, in which the particles are moved iteratively from samples from the prior to samples from the posterior, such as in iterative Ensemble Kalman Filters or Smoothers (e.g., the discussion in van Leeuwen and Evensen, 1996, Gu and Oliver, 2007, Emerick and Reynolds, 2013, Asch *et al.*, 2016),

or explore transportation particle filters (e.g., Villani 2008, Reich, 2011, Liu and Wang, 2016, Pulido and van Leeuwen, 2019), or

4. Move the particles from prior to posterior in one step (e.g., Moselhy and Marzouk, 2012, Ades and van Leeuwen, 2013, Zhu *et al.*, 2016), and if the posterior is Gaussian use a deterministic EnKF (e.g., Bishop *et al.*, 2001, Anderson, 2001, Whitaker and Hamill, 2002, or
5. Use Markov-chain Monte-Carlo methods to generate samples from the posterior in a sequential way (e.g., Robert and Casella, 2004), or
6. Draw directly from the posterior. This can rarely be done, but a stochastic EnKF is such a scheme.

As shown in the main text, deterministic ensemble Kalman Filters belong to method 4, and the stochastic version of the EnKF belongs to method 6.

## B. Posterior covariance via Bayes' Theorem

Instead of using the direct derivation of the posterior covariance from Bayes' Theorem via multiplication of the prior and the likelihood and completing the squares under the exponent, in this Appendix we demonstrate a derivation based on the expectation operator. This will allow us to clearly see the treatment of the observations in the derivation.

To keep track of the random variables, we use the notation  $E_a[...]$  to denote an expectation with respect to the pdf of  $\mathbf{a}$ . We find, using  $\mathbf{x}^a = (\mathbf{I} - \mathbf{KH})\mathbf{x}^f + \mathbf{K}\mathbf{y}^o$ , and realizing that the expectation should be taken with the posterior pdf in Bayes' Theorem:

$$\begin{aligned} \mathbf{P} &= E_{\mathbf{x}|\mathbf{y}^o} [(\mathbf{x} - \mathbf{x}^a)(\mathbf{x} - \mathbf{x}^a)^T] \\ &= E_{\mathbf{x}|\mathbf{y}^o} [(\mathbf{I} - \mathbf{KH})(\mathbf{x} - \mathbf{x}^f) - \mathbf{K}(\mathbf{y}^o - \mathbf{H}\mathbf{x})] \\ &\quad \times [(\mathbf{I} - \mathbf{KH})(\mathbf{x} - \mathbf{x}^f) - \mathbf{K}(\mathbf{y}^o - \mathbf{H}\mathbf{x})]^T \\ &= (\mathbf{I} - \mathbf{KH})E_{\mathbf{x}|\mathbf{y}^o} [(\mathbf{x} - \mathbf{x}^f)(\mathbf{x} - \mathbf{x}^f)^T] (\mathbf{I} - \mathbf{KH})^T \\ &\quad + E_{\mathbf{x}|\mathbf{y}^o} [\{ \mathbf{K}(\mathbf{H}\mathbf{x} - \mathbf{y}^o) \} \{ \mathbf{K}(\mathbf{H}\mathbf{x} - \mathbf{y}^o) \}^T] \\ &\quad - (\mathbf{I} - \mathbf{KH})E_{\mathbf{x}|\mathbf{y}^o} [(\mathbf{x} - \mathbf{x}^f) \{ \mathbf{K}(\mathbf{y}^o - \mathbf{H}\mathbf{x}) \}^T] \\ &\quad - E_{\mathbf{x}|\mathbf{y}^o} [\{ \mathbf{K}(\mathbf{y}^o - \mathbf{H}\mathbf{x}) \} (\mathbf{x} - \mathbf{x}^f)^T] (\mathbf{I} - \mathbf{KH})^T. \quad (\text{B1}) \end{aligned}$$

To evaluate these expectations, we need to be mindful that the pdf in the integrals is the posterior pdf. For the first term we proceed as follows, using  $(\mathbf{x} - \mathbf{x}^f)p(\mathbf{x}) = -\mathbf{B}\partial p(\mathbf{x})/\partial \mathbf{x}$ :

$$\begin{aligned} E_{\mathbf{x}|\mathbf{y}^o} [(\mathbf{x} - \mathbf{x}^f)(\mathbf{x} - \mathbf{x}^f)^T] &= \int (\mathbf{x} - \mathbf{x}^f)(\mathbf{x} - \mathbf{x}^f)^T p(\mathbf{x}|\mathbf{y}^o) d\mathbf{x} \\ &= - \int \mathbf{B} \frac{\partial p(\mathbf{x})}{\partial \mathbf{x}} (\mathbf{x} - \mathbf{x}^f)^T \frac{p(\mathbf{y}^o|\mathbf{x})}{p(\mathbf{y}^o)} d\mathbf{x} \end{aligned}$$

$$\begin{aligned} &= \mathbf{B} + \int \mathbf{B} \frac{\partial p(\mathbf{y}^o|\mathbf{x})}{\partial \mathbf{x}} (\mathbf{x} - \mathbf{x}^f)^T \frac{p(\mathbf{x})}{p(\mathbf{y}^o)} d\mathbf{x} \\ &= \mathbf{B} + \int \mathbf{B} \{ \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{y}^o - \mathbf{H}\mathbf{x}) \} (\mathbf{x} - \mathbf{x}^f)^T p(\mathbf{x}|\mathbf{y}^o) d\mathbf{x} \\ &= \mathbf{B} + E_{\mathbf{x}|\mathbf{y}^o} [\mathbf{B} \{ \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{y}^o - \mathbf{H}\mathbf{x}) \} (\mathbf{x} - \mathbf{x}^f)^T], \quad (\text{B2}) \end{aligned}$$

where we used partial integration and

$$\partial p(\mathbf{y}^o|\mathbf{x})/\partial \mathbf{x} = \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{y}^o - \mathbf{H}\mathbf{x}) p(\mathbf{y}^o|\mathbf{x}).$$

The first term in Equation (B1) now becomes, exploring the identity  $(\mathbf{I} - \mathbf{KH})\mathbf{B}\mathbf{H}^T \mathbf{R}^{-1} = \mathbf{K}$ ,

$$\begin{aligned} (\mathbf{I} - \mathbf{KH})E_{\mathbf{x}|\mathbf{y}^o} [(\mathbf{x} - \mathbf{x}^f)(\mathbf{x} - \mathbf{x}^f)^T] (\mathbf{I} - \mathbf{KH})^T &= (\mathbf{I} - \mathbf{KH})\mathbf{B}(\mathbf{I} - \mathbf{KH})^T \\ &\quad + E_{\mathbf{x}|\mathbf{y}^o} [\{ \mathbf{K}(\mathbf{y}^o - \mathbf{H}\mathbf{x}) \} (\mathbf{x} - \mathbf{x}^f)^T] (\mathbf{I} - \mathbf{KH})^T. \quad (\text{B3}) \end{aligned}$$

Similarly, we find

$$\begin{aligned} E_{\mathbf{x}|\mathbf{y}^o} [\{ \mathbf{K}(\mathbf{H}\mathbf{x} - \mathbf{y}^o) \} \{ \mathbf{K}(\mathbf{H}\mathbf{x} - \mathbf{y}^o) \}^T] &= \int \{ \mathbf{K}(\mathbf{H}\mathbf{x} - \mathbf{y}^o) \} \{ \mathbf{K}(\mathbf{H}\mathbf{x} - \mathbf{y}^o) \}^T p(\mathbf{x}|\mathbf{y}^o) d\mathbf{x} \\ &= \int (\mathbf{I} - \mathbf{KH})\mathbf{B} \frac{\partial p(\mathbf{y}^o|\mathbf{x})}{\partial \mathbf{x}} \{ \mathbf{K}(\mathbf{y}^o - \mathbf{H}\mathbf{x}) \}^T \frac{p(\mathbf{x})}{p(\mathbf{y}^o)} d\mathbf{x} \\ &= (\mathbf{I} - \mathbf{KH})\mathbf{B}\mathbf{H}^T \mathbf{K}^T \\ &\quad + \int (\mathbf{I} - \mathbf{KH})(\mathbf{x} - \mathbf{x}^f) \{ \mathbf{K}(\mathbf{y}^o - \mathbf{H}\mathbf{x}) \}^T p(\mathbf{x}|\mathbf{y}^o) d\mathbf{x} \\ &= (\mathbf{I} - \mathbf{KH})\mathbf{B}\mathbf{H}^T \mathbf{K}^T \\ &\quad + (\mathbf{I} - \mathbf{KH})E_{\mathbf{x}|\mathbf{y}^o} [(\mathbf{x} - \mathbf{x}^f) \{ \mathbf{K}(\mathbf{y}^o - \mathbf{H}\mathbf{x}) \}^T]. \quad (\text{B4}) \end{aligned}$$

Combining these two results in Equation (B1), we obtain, using  $(\mathbf{I} - \mathbf{KH})\mathbf{B}\mathbf{H}^T \mathbf{K} = \mathbf{K}\mathbf{R}\mathbf{K}^T$ ,

$$\mathbf{P} = (\mathbf{I} - \mathbf{KH})\mathbf{B}(\mathbf{I} - \mathbf{KH})^T + \mathbf{K}\mathbf{R}\mathbf{K}^T, \quad (\text{B5})$$

because all remaining integral terms cancel each other. The important point of this derivation is that observations are indeed treated as constants, not as random variables because there is no integration over the observations. Hence perturbing observations is not needed to derive the posterior covariance, and indeed Bayes' Theorem tells us that we should not do that in the first place.

## C. Posterior covariance via the BLUE

Note that in this Appendix  $\mathbf{y}$  is a random variable. We derive the Joseph form of the posterior covariance using the BLUE. This can be found in textbooks but the explicit

derivation will allow us to understand how observations are treated in a stochastic EnKF.

The posterior covariance is defined in the BLUE as

$$\begin{aligned} \mathbf{P} &= \int (\mathbf{x} - \mathbf{x}^a)(\mathbf{x} - \mathbf{x}^a)^T p(\mathbf{x}, \mathbf{y}) \, d\mathbf{x} \, d\mathbf{y} \\ &= E_{\mathbf{x}, \mathbf{y}} [(\mathbf{x} - \mathbf{x}^a)(\mathbf{x} - \mathbf{x}^a)^T]. \end{aligned} \quad (\text{C1})$$

Note the use of the joint pdf in the expectation. We can now proceed as in Appendix B, using  $\mathbf{x}^a = (\mathbf{I} - \mathbf{KH})\mathbf{x}^f + \mathbf{Ky}$ , which is true for any value for the observation:

$$\begin{aligned} \mathbf{P} &= (\mathbf{I} - \mathbf{KH})E_{\mathbf{x}, \mathbf{y}} [(\mathbf{x} - \mathbf{x}^f)(\mathbf{x} - \mathbf{x}^f)^T] (\mathbf{I} - \mathbf{KH})^T \\ &\quad + E_{\mathbf{x}, \mathbf{y}} [\{\mathbf{K}(\mathbf{H}\mathbf{x} - \mathbf{y})\} \{\mathbf{K}(\mathbf{H}\mathbf{x} - \mathbf{y})\}^T] \\ &\quad - (\mathbf{I} - \mathbf{KH})E_{\mathbf{x}, \mathbf{y}} [(\mathbf{x} - \mathbf{x}^f)\{\mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x})\}^T] \\ &\quad - E_{\mathbf{x}, \mathbf{y}} [\{\mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x})\} (\mathbf{x} - \mathbf{x}^f)^T] (\mathbf{I} - \mathbf{KH}). \end{aligned} \quad (\text{C2})$$

In this case, one can split integrals over  $\mathbf{x}$  and  $\mathbf{y}$  to find:

$$\begin{aligned} &E_{\mathbf{x}, \mathbf{y}} [(\mathbf{x} - \mathbf{x}^f)(\mathbf{x} - \mathbf{x}^f)^T] \\ &= \int \int (\mathbf{x} - \mathbf{x}^f)(\mathbf{x} - \mathbf{x}^f)^T p(\mathbf{x}, \mathbf{y}) \, d\mathbf{x} \, d\mathbf{y} = \mathbf{B} \end{aligned} \quad (\text{C3})$$

by integrating the  $\mathbf{y}$  dependence out. Furthermore, for the second term in the expression for  $\mathbf{P}$  we find:

$$\begin{aligned} &E_{\mathbf{x}, \mathbf{y}} [\{\mathbf{K}(\mathbf{H}\mathbf{x} - \mathbf{y})\} \{\mathbf{K}(\mathbf{H}\mathbf{x} - \mathbf{y})\}^T] \\ &= \int \int \mathbf{K}(\mathbf{H}\mathbf{x} - \mathbf{y})(\mathbf{H}\mathbf{x} - \mathbf{y})^T \mathbf{K}^T p(\mathbf{x}, \mathbf{y}) \, d\mathbf{x} \, d\mathbf{y} \\ &= \int \left[ \int \mathbf{K}(\mathbf{H}\mathbf{x} - \mathbf{y})(\mathbf{H}\mathbf{x} - \mathbf{y})^T \mathbf{K}^T p(\mathbf{y}|\mathbf{x}) \, d\mathbf{y} \right] p(\mathbf{x}) \, d\mathbf{x} \\ &= \int \left[ \int \mathbf{K}(-\epsilon)(-\epsilon)^T \mathbf{K}^T p(\epsilon) \, d\epsilon \right] p(\mathbf{x}) \, d\mathbf{x} \\ &= \int \mathbf{K} \mathbf{R} \mathbf{K}^T p(\mathbf{x}) \, d\mathbf{x} = \mathbf{K} \mathbf{R} \mathbf{K}^T, \end{aligned} \quad (\text{C4})$$

where we used that  $\mathbf{y} = \mathbf{H}\mathbf{x} + \epsilon$  when  $\mathbf{x}$  is given, allowing for the change of variables from  $\mathbf{y}$  to  $\epsilon$ . It is easy to see that the cross terms are zero, leading to the correct expression for  $\mathbf{P}$ . One might expect that this derivation leads to observation perturbations when evaluated via a Monte-Carlo sampling, but the main text shows that this is not the case.